THE INCIDENCE OF A DISEASE IN POPULATION GROUPS, THE NUMBER OF PEOPLE IN WHICH IS KNOWN OR UNKNOWN

BY

J. W. W. STEPHENS

(*Received for publication* 25 *May*, 1922)

As an example of the 'incidence,' 'occurrence,' or 'distribution' of cases of a disease in one or more groups, such as age-groups of a population, the number of people in which is *unknown*, we may take the following. Of a total of twenty cases of influenza, let us suppose that ten occurred in Group A and ten in Group B, then the respective incidences ten and ten are equal, and the number that occur in each group per one hundred cases, viz., fifty and fifty, are also equal.

As an example of the 'incidence,' 'occurrence' or 'distribution' in age-groups, the number of people in which is *known*, we may take the following. Of a total of twenty cases of influenza, let us suppose that ten occurred in Group A, containing one hundred people, and ten in Group B, containing fifty people, then the incidences are 10 per cent. and 20 per cent. respectively (and the ratios of the incidences per one hundred cases 33 per cent. and 66 per cent. respectively).

It will be evident that the term 'incidence' has been used here in two different senses. In the first sense of the term, 'incidence,' it is only the number of cases that is known. In the second sense, when not only the number of cases but also the number of people among whom the cases occur is known, the term is applied to a figure expressing the number of cases that occur per one hundred people in each group.

To emphasise the distinction in meaning between these two uses of the term 'incidence,' it would seem advisable to confine the term 'incidence' to the use of the term in the first sense, and the term 'incidence rate' to the use of the term in the second sense.

In practice, however, certain deductions are often made when the number of cases alone is known, which can, as we shall see, be only justifiably made when the number of people in the groups is also known, *i.e.*, when the '*incidence rate*' can be calculated.

In regard to 'incidence,' the larger the group, the larger (*ceteris paribus*) is the incidence. In regard to 'incidence rates,' the factor of unequal size of the groups, if it exists, is eliminated, as the rate is calculated for one hundred people in each group.

The above examples may be tabulated as follows, using the words incidence and incidence rates in the sense defined above.

TABLE I.

Shewing distinction between incidence and incidence rate.

		1	2	3	4	5	6	7
	Group	Total number of people that occur in each group (Census)	Number of people that occur in each group per 100 people (Census)	Incidence, *i.e.*, total number of cases observed that occur in each group	Number of cases that occur in each group per 100 cases	Incidence rate, *i.e.*, number of cases occurring among 100 people in each group	Ratios of the incidence rates to one another	Ratios of the incidence rates to one another per cent.
Ex. 1.	A	...	...	10	50	...	...	...
	B	...	...	10	50	...	...	...
Ex. 2.	A	100	66·6	10	50	10	1	33·3
	B	50	33·3	10	50	20	2	66·6

From the second example in the table we see that the 'liability to attack' of a person in Group A is 10 per cent. and in Group B 20 per cent., *i.e.*, it is twice as great in Group B as in Group A. This fact cannot, however, be deduced from the figures in the first example, because, although the number of cases is the same as in the second example, nothing is known as to the number of people among whom the cases occurred. It is the incidence rates (actual or relative) that are of importance if we are studying what may be termed the 'real incidence' of the disease on a group.

DIABETES

We find recorded in Osler and Macrea, *System of Medicine*, second edition, p. 675, the age-group incidence of three hundred and thirty-five cases of diabetes in Baltimore (column 3) from which can be readily calculated the age-group incidences per one hundred cases (column 4). The figures for the age-group, distribution or incidence of the population of Baltimore per one hundred people are not given, so that for purposes of illustration I have used those of Liverpool as deduced from the 1911 census (Table II, column 2).

TABLE II.

Showing incidences and ratios of incidence rates in Diabetes.

Age-Group	1 Total number of people that occur in each group (Census)	2 Number of people that occur in each group per 100 people (Census)	3 Incidence, *i.e.*, total number of cases observed that occur in each group	4 Number of cases that occur in each group per 100 cases	5 Incidence rate, *i.e.*, number of cases occurring among 100 people in each group	6 Ratios of the incidence rates to one another	7 Ratios of the incidence rates to one another per cent.
1–10 ...	...	23·3	8	2·18	...	0·0935	0·86
11–20 ...	...	18·9	25	7·34	...	0·3883	3·59
21–30 ...	...	16·7	44	13·1	...	0·7245	6·71
31–40 ...	...	15·8	61	18·2	...	1·1519	10·67
41–50 ...	...	11·2	69	20·6	...	1·8392	17·04
51–60 ...	...	7·3	89	26·5	...	3·6301	33·64
61–70 ...	...	4·5	33	9·8	...	2·1111	19·47
71–80 ...	...	2·0	6	1·7	...	0·8500	7·87
81– ...	...	0·3	0	0·0	...	0·0	0·0
	...	100·0	335	99·42	...	10·7886	99·85

If we were dealing with the total number of people in each group (column 1) instead of the number per one hundred of the population, and divided a figure in column 3 by the corresponding figure in column 1 and multiplied the result by one hundred, the figures obtained would represent the incidence rates, *i.e.*, the incidence per

one hundred people (column 5). But, in the present case, where we have divided the percentage figures in column 4 by the percentage figures in column 2, the resulting figures (column 6) represent simply the *ratios** which the incidence rates bear to one another, and from these we can easily calculate the ratios, when the sum of the ratios is one hundred (column 7). Thus, to refer to Table II (column 7), we see that of one hundred cases of diabetes about thirty-three would occur among so many people in the age-group 51-60, while about half that number (17·04) would occur among the *same number* of people in the age-group 41-50, whereas, considering the incidence merely (column 4), it is about the same in the two groups, viz., 20·6 and 26·5 respectively.

INFLUENZA

The following example (Table III) is taken from Nothnagel's *Encyclopædia of Practical Medicine*, English Edition. Article 'Influenza,' p. 571. The actual figures for the case incidence and the population incidence in the various groups are not given, but only the percentage incidences in each case, in the form of graphs. The figures are only approximately correct, as it was not possible to calculate them exactly from the graphs. As in Table II, by dividing the percentages in column 4 by the corresponding ones in column 2, we get a series of figures (column 6) which represent the ratios which the incidence rates bear to one another, and in column 7 the ratios of these rates per cent. Thus, the 'liability to attack' (column 7) in the age period 21-30 is slightly more than twice as great as in the age-period 51-60, but what the actual figures for liability to attack are it is impossible to say, as it is only the percentage and not the actual number of people in the groups that is known. The figures in column 7 are not comparable with those in column 4; strictly speaking, no conclusions as to 'liability to attack' can be based on the figures in column 4 by themselves. It is only if we assume *some* knowledge of the number of people in the groups that the case incidence figures have any value in this respect.

*The *ratios*, but of course not the same actual figures, in this column could equally well be got by dividing the figures in column 3 by those in column 2.

Thus, we could probably infer that the liability to attack was greater in the 21-30 period than in the 11-20 period, because we *assume* that the population of the 21-30 period is probably not twice that of the

TABLE III.

Showing incidences and ratios of incidence rates in Influenza.

	1	2	3	4	5	6	7
Age-Group	Total number of people that occur in each group (Census)	Number of people that occur in each group per 100 people (Census)	Incidence, *i.e.*, total number of cases observed that occur in each group	Number of cases that occur in each group per 100 cases	Incidence rate, *i.e.*, number of cases occurring among 100 people in each group	Ratios of the incidence rates to one another	Ratios of the incidence rates to one another per cent.
1–10 ...	...	19	...	8	...	0·421	5·86
11–20 ...	...	16	...	15	...	0·937	13·05
21–30 ...	...	21	...	32	...	1·523	21·07
31–40 ...	...	16	...	20	...	1·250	17·40
41–50 ...	...	12	...	14	...	1·166	16·23
51–60 ...	...	8	...	6	...	0·750	10·44
61–70 ...	...	5	...	4	...	0·800	11·14
71–80 ...	...	3	...	1	...	0·333	4·63
	...	100	...	100	...	7·180	99·82

11-20 period; but we can only make accurate deductions, giving relative or actual figures, when we base them on the number of people, relative or actual, in the groups.

BLACKWATER FEVER

It has been commonly stated that the liability to an attack of blackwater fever is greater in persons infected with malignant tertian parasites than in those infected with simple tertian or quartan parasites. These statements are based on the particular parasites present in so many *cases* of blackwater fever, but, as we have shown

above, no conclusions can be drawn as to liability to attack unless we have population data as well.

The case before us is parallel with the two examples we have already considered, though here, instead of age-groups, we have groups of persons (malaria cases) infected with the malignant tertian and simple tertian parasites respectively (Table IV). The data are

TABLE IV.

Showing relative liability to an attack of Blackwater fever of persons infected with malignant tertian and simple tertian parasites respectively.

		1	2	3	4	5	6	7
	Parasite Group	Total number of case of Malaria that occur in each group	Number of cases of Malaria that occur in each group per 100 cases of Malaria	Total number of Blackwater cases observed that occur in each group	Number of cases of Blackwater that occur in each group per 100 cases of Blackwater	Incidence rate, *i.e.*, number of cases of Blackwater that occur among 100 cases of Malaria in each group	Ratios of incidence rates to one another	Ratios of incidence rates per cent.
Ex. 1	Malignant tertian ...	...	74	...	76·4	...	1·032	53·2
	Simple tertian ...	...	26	...	23·6	...	0·908	46·8
		...	100	...	100	...	1·94	100·0
Ex. 2	Malignant tertian ...	...	68·45	...	54·03	...	0·789	35·1
	Simple tertian ...	...	31·56	...	45·96	...	1·456	64·9
		...	100·00	...	99·99	...	2·245	99·9

taken from a paper in the *Annals of Tropical Medicine and Parasitology*, Vol. VII, December, 1913, p. 487, in which I have summarised the data of Deeks and James, and Lovelace, respectively.

As before, the figures in column 6 are got by dividing those in column 4 by the corresponding ones in column 2. The figures in column 7 are then calculated for one hundred cases.

In the first example, the incidence rate of blackwater fever in malignant tertian infections is only slightly greater than that in simple tertian infections.

In the second example, the incidence rate in simple tertian infections is nearly twice as great as that in malignant tertian infections.

We are not concerned here with the discrepancy between the results, but with the fact that in each case deductions based solely on the incidence, *i.e.*, occurrence of malignant tertian or simple tertian parasites in the blackwater *cases*, would have led to different but erroneous conclusions.

A reference to the current text-books of Tropical Medicine would afford many other examples of a similar kind, where conclusions are drawn from a knowledge of the number of cases only, in the absence of any knowledge of the number of people among whom the cases occur.